AD-A283 382

DOT/FAA/CT-94/87

FAA Technical Center
Atlantic City International Airport,
N.J. 08405

Test and Evaluation Plan (TEP) for Improvised Explosive Device Sceening System (IEDSS)

J. L. Fobes, Ph. D.
Ronald J. Lofaro, Ph. D.

September 1994

Project Plan

U.S. Department of Transportation
Federal Aviation Administration

94-35426

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof. The United States Government does not endorse products or manufacturers. Trade or manufacturer's names appear herein solely because they are considered essential to the object of this report.

1 Report No	2 Government Accession No	3 Recipient's Catalog No
DOT/FAA/CT-94/87		

4 Title and Subtitle	5 Report Date
TEST AND EVALUATION PLAN (TEP) FOR IMPROVISED EXPLOSIVE DEVICE SCREENING SYSTEMS (IEDSS)	August 1994
	6 Performing Organization Code
	ACA-400

7 Author(s)	8 Performing Organization Report No
James L. Fobes, Ph.D., Ronald J. Lofaro, Ph.D., (Federal Aviation Administration) Robert L. Malone, Douglas C. Fischer, Jack Berkowitz. (Galaxy Scientific Corporation)	

9 Performing Organization Name and Address	10 Work Unit No (TRAIS)
Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405	
	11 Contract or Grant No
	DTFA03-89-C-00043

12 Sponsoring Agency Name and Address	13 Type of Report and Period Covered
U.S. Department of Transportation Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405	Draft
	14 Sponsoring Agency Code
	ACS-1

15 Supplementary Notes

16 Abstract

This document is the Operational Test and Evaluation Plan (OT&E) to evaluate the effectiveness of X-ray baggage screening equipment for Improvised Explosive Devices (IEDs) in cluttered baggage.

The Improvised Explosive Device Screening Systems (IEDSS) to be tested during the OT&E consist of standard (black and white) and enhanced ("color") display X-ray equipment and their operators. The X-ray equipment will be used to scan for simulated IEDs in both checked (C) and carry-on (CO) cluttered baggage.

The OT&E will be conducted at San Francisco International Airport (SFO). The results will be analyzed and become part of a later document.

17 Key Words	18 Distribution Statement
X-ray X-ray screening signal detection explosive device screening systems	This document is available through the National Technical Information Service (NTIS), Springfield, Virginia 22161

19 Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F1700.7 (8-72)   Reproduction of completed page authorized

OPERATIONAL

TEST AND EVALUATION PLAN

(TEP)

IMPROVISED EXPLOSIVE DEVICE SCREENING SYSTEMS

(IEDSS)

Prepared for

FEDERAL AVIATION ADMINISTRATION
AVIATION SECURITY HUMAN FACTORS PROGRAM
AVIATION SECURITY R&D SERVICE
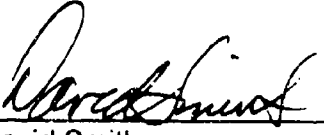FAA TECHNICAL CENTER
ACA-400

Approved:

J. L. Fobes, Ph.D.
Program Manager
Aviation Security Human Factors Program

A-1

and

CIVIL AVIATION SECURITY - OFFICE OF POLICY & PLANNING
FAA WASHINGTON HEADQUARTERS

Reviewed:

James Farrell
Task Force Manager
ACS EDS Task Force

David Smith
Acting Manager
Technology Integration Division
ACP-400

PREFACE

This test plan was developed to meet the Critical Operational Issues and Criteria (COIC) set forth by the Federal Aviation Administration (FAA). The key FAA personnel supporting this testing effort were James L. Fobes, Ph.D. and Ronald Lofaro, Ph.D., both engineering Research Psychologists of the Aviation Security Research and Development Service at the FAA Technical Center (FAATC).

Galaxy Scientific Corporation (GSC) prepared this document under contract number DTFA03-89-C-00043 with the FAATC. The Program Manager at GSC is William Hassler, Jr. The authors of this document are Robert Malone, Doug Fischer, and Jack Berkowitz.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF SYMBOLS AND ACRONYMS

C	Checked
CO	Carry-on
COIC	Critical Operational Issues and Criteria
FAA	Federal Aviation Administration
IED	Improvised Explosive Device
IEDSS	Improvised Explosive Device Screening Systems
MBS	Modular Bomb Set
MNS	Mission Needs Statement
MOP	Measures of Performance
NAS	National Airspace System
OT&E	Operational Test and Evaluation
Pd	Probability of Detection
Pfa	Probability of False Alarm
RPI	Research Project Initiative
SDT	Signal Detection Theory
SFO	San Francisco International Airport
TEP	Test and Evaluation Plan
TER	Test and Evaluation Report
TnT	Training and Testing
UA	United Airlines

1.   INTRODUCTION.

1.1   PURPOSE.

While new systems and methods are being developed to detect Improvised Explosive Devices (IEDs) and other hazardous objects, no industry or Federal Aviation Administration (FAA) standards or guidelines exist to optimize IED detection performance with current or future x-ray equipment and operators. This Test and Evaluation Plan (TEP) supports the FAA's independent evaluation of the effectiveness of x-ray baggage screening for detecting IEDs in cluttered baggage. It also supports an evaluation of x-ray screener performance enhancements through IED detection training.

1.2   SCOPE.

The Improvised Explosive Device Screening Systems (IEDSS) to be tested during the Operational Test and Evaluation (OT&E) consist of standard (black/white) and enhanced ("color") display x-ray equipment and their operators. The x-ray equipment will be used to scan for simulated IEDs in both checked (C) and carry-on (CO) cluttered baggage, as shown in figure 1. This OT&E will focus on determining each system's operational effectiveness in meeting the requirements set forth in the Critical Operational Issues and Criteria (COIC), contained in section 1.6 of this TEP. The CO results will also be evaluated by comparisons with IED test-kit detection data previously collected by the FAA with uncluttered CO bags.
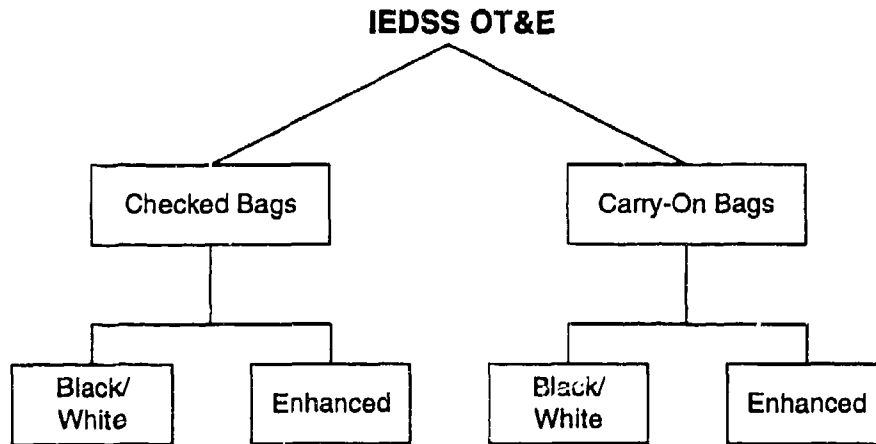
FIGURE 1. IEDSS TO BE EVALUATED

1.3   BACKGROUND.

The threat to civil aviation security has changed dramatically in the last decade. In the 1980s, the threat was hijacking; the FAA's role in aviation security was greatly expanded during this period, especially after the 1985 hijacking of TWA Flight 847 in the Middle East. In the 1990s, there has been a shift from hijacking to the threat of sabotage by bombings.

Improvements in technology available to hostile elements, especially in the area of explosive devices, have resulted in an increased airliner vulnerability. Terrorists are reducing their use of prefabricated

explosive devices, such as grenades, and opting for less detectable IEDs. An IED can be made from a variety of materials that may resemble "innocent" or everyday objects, such as batteries, wires, and digital clocks. For example, plastic explosives made with Semtex and C-4 can be shaped and molded into sheets or blocks that, when passed through x-ray screening devices, appear as innocent items such as books or radios. Terrorists have also learned to embed IEDs in electronic devices, making detection even more difficult, as in the Pan American Flight 103 disaster. In addition, miniaturization and digitization of timing devices compound the problem of IED detection with x-ray screening.

Because the focus of civil aviation security has shifted from hijackings to bombings, the need for improvements in x-ray screener equipment, operator training, and overall system performance has increased. Sophisticated and dedicated terrorists possess the knowledge, materials, and capability to build difficult-to-detect IEDs. The potential for complete aircraft destruction, with hundreds of fatalities and the disruption of the National Airspace System (NAS), has increased.

The ACS and ACP headquarters elements of the FAA have identified the need for research into the performance of IEDSS (and particularly the human component) in detecting IEDs in C and CO baggage. This research will focus on comparing operator performance in detecting IEDs, with black/white and enhanced x-ray equipment, under operational conditions although the screener will know when testing starts and ends. Since baggage screeners are now the main line of defense for detecting IEDs to ensure aviation security, it is essential to evaluate the effectiveness of IEDSS.

This research is being conducted under the FAA's Aviation Security Human Factors Program, Research Project Initiative (RPI) 127 in support of Mission Needs Statement (MNS) 163.

1.4   SYSTEM DESCRIPTION.

The OT&E will evaluate overall system performance with black/white and enhanced x-ray equipment.

1.4.1   Black/White X-Ray Equipment.

Black/white x-ray equipment transmits an x-ray beam through baggage whose contents absorb x-rays differentially. As determined by baggage and content densities, a resulting black/white image is displayed to baggage screener personnel. This system relies on the screener to detect patterns characteristic of IEDs and other threatening materials.

1.4.2   Enhanced X-Ray Equipment.

The EG&G E-Scan system has dual energy and color image features. It uses color to depict the image as organic (light elements), inorganic (usually heavy elements), and opaque materials (a great deal of heavy element matter). For instance, the E-Scan system assigns the color orange to organic materials, which might include explosives. Besides the pattern cues presented in black/white displays, enhanced x-ray equipment assists screeners in detecting IEDs by displaying organic-based color cues and provides screeners with the capability to view only the organic elements of baggage contents.

1.5 COMPUTER-BASED INSTRUCTION TRAINING.

EG&G's Training and Testing (TnT) system delivers computer-based training modules on x-ray image identification of IEDs with black/white and enhanced displays. The instruction modules to be used in the test are self-paced and contain instructions and formal training to be given to the screeners.

1.6 CRITICAL OPERATIONAL ISSUES AND CRITERIA.

The following operational issues will be tested and evaluated. Their scope, criteria, rationale, evaluation approach, Measures of Performance (MOPs), analysis methodology, and data presentations are described in Chapter 2.

1.6.1 Issue 1.

Can baggage screeners detect IEDs using black/white x-ray equipment?

1.6.2 Issue 2.

Can IED detectability with black/white x-ray equipment be improved with a training intervention?

1.6.3 Issue 3.

Can baggage screeners better detect IEDs using enhanced x-ray equipment?

1.6.4 Issue 4.

Can IED detectability with enhanced x-ray equipment be improved with a training intervention?

1.7   TEST AND EVALUATION MILESTONES.

Table 1 shows the milestones that have been established to ensure orderly execution of the test and evaluation processes for planning, programming, and reporting.

TABLE 1. TEST AND EVALUATION MILESTONES

Event Completion Schedule	
Prepare TEP	Apr. 1994
Coordinate Test Site	May 1994
Obtain IED Test Objects and Baggage	May 1994
Coordinate Screeners' Assignments	May 1994
Conduct Pilot Test	Jun. 1994
Conduct Operational Test	Jul.–Aug. 1994
Present Executive Summary and Major Quantitative Findings to Task Force	Aug. 1994
Prepare Test and Evaluation Report (TER) and Present TER to Task Force	Sep. 1994

2.   TEST AND EVALUATION STRATEGY.

2.1   EVALUATION CONCEPT.

The IEDSS effectiveness will be evaluated against the COIC. The primary measures are the probability of detection (Pd) and the probability of false alarm (P_{fa}) for Modular Bomb Set (MBS)-simulated IEDs using black/white or enhanced x-ray equipment. Comparisons will also be made with both C and CO passenger bags using the two x-ray technologies. CO data will further be contrasted with that recently collected by the FAA using three of the same MBS configurations in a non-cluttered CO bag.

2.1.1   Evaluation Approach.

a.   The OT&E protocol will be conducted at the SFO Category X airport. Before the OT&E, contractor and FAA personnel will travel to the test site to coordinate operations with essential airport, security, airline, and law enforcement personnel. These trips will also be used to ensure that test and evaluation personnel become familiar with the baggage screening environments.

b.   To establish the appropriate procedures and minimize problems in the OT&E, a pilot study will be conducted. It will determine if data collection procedures interfere with baggage screeners' tasks and whether established OT&E procedures are appropriate. The contractor will prepare a report documenting difficulties encountered during the pilot study and suggest modifications to the protocol to continue the IEDSS OT&E.

2.1.2   Experimental Design.

a.   The experimental approach will be a mixed-factorial design consisting of two phases, Pre-Training and Post-Training, which surround computer-based IED detection training. There will

4

be four groups within each phase: C-black/white, C-enhanced, CO-black/white, and CO-enhanced. There will be a total of 40 subjects participating in the OT&E with 10 subjects per group.

b. During the OT&E, each C bag screener will view 220 Pre-Training and 220 Post-Training bags for their assigned IEDSS configuration (black/white or enhanced). Thus, each screener will view a total of 440 bags, of which 40 (20 Pre-Training and 20 Post-Training) will contain the MBS-simulated IEDs.

c. During the OT&E, each CO bag screener will view 132 Pre-Training and 132 Post-Training bags for their assigned equipment condition (black/white or enhanced). Thus, each screener will view a total of 264 bags, of which 24 (12 Pre-Training and 12 Post-Training) will contain the MBS-simulated IEDs.

d. The Pre-Training phase will establish a detection performance score baseline before IED detection training. The Post-Training phase will be compared to the Pre-Training to assess the effectiveness of the computer-based training. The computer-based training will be given to the baggage screeners in each group. The order of inserting the MBS-configured baggage into the standard passenger bag flow will be randomized. Table 2 shows the experimental design for this study.

TABLE 2. EXPERIMENTAL DESIGN USED IN THIS STUDY

		Within Group		
Between Group		Pre-Training	Computer-Based Training	Post-Training
Checked: B/W Enhanced				
Carry-On: B/W Enhanced				

e. Overall system effectiveness will be analyzed by determining the extent to which baggage screeners can successfully identify MBSs using current x-ray equipment. Systems will be evaluated in terms of MBS P_d and in terms of a Signal Detection Theory (SDT) paradigm. Each COIC will be analyzed using a combination of statistical techniques.

2.1.3   Significant Test and Evaluation Limitations.

None identified at this time.

2.2   OPERATIONAL EFFECTIVENESS.

The operational effectiveness of IEDSS will be addressed through COIC concerning system performance with C and CO bags. Pre- and Post-Training conditions will be administered to subjects using both black/white and enhanced x-ray equipment.

2.2.1 Issue 1. Black/White X-Ray Detection.

Can baggage screeners detect IEDs using black/white x-ray equipment?

2.2.1.1 Scope.

This issue examines the effectiveness of baggage screeners using black/white x-ray based equipment to detect MBSs in both C and CO cluttered baggage. 'Representative' passenger bags containing FAA MBS test objects will be inserted into the normal baggage screening operations. Representative refers to baggage size, density, clutter, and organic content.

2.2.1.2 Criteria.

None. This issue is investigative in nature.

2.2.1.3 Rationale.

Black/white x-ray based IEDSS are common in U.S. airports, but have an unknown effectiveness when used for detecting IEDs in cluttered bags. Inserting MBS test objects within cluttered bags provides crucial information. This issue is investigative in nature because ACA does not have a basis for determining the minimal acceptable P_d with black/white x-ray equipment.

2.2.1.4 Evaluation Approach.

The P_d and SDT paradigm measures will be used to assess system performance in the operational environment.

2.2.1.5 MOPs.

MOP 1. The P_d and P_{fa} of various MBS configurations in C bags with black/white x-ray.
MOP 2. The P_d and P_{fa} of various MBS configurations in CO bags with black/white x-ray.

2.2.1.6 Analysis Methodology and Data Presentations.

a.   Detection rates will be tabulated in a .. by 2 signal detection matrix and grouped according to C and CO baggage, as shown in figure 2.

b. P_d and P_{fa} will be determined across groups.

c. β and d' values will be calculated for C and CO bags.

State of MBS Image

	MBS Present	MBS Not Present
Yes	Hit	False Alarm
Test Subject Response		
No	Miss	Correct Rejection

FIGURE 2. 2 BY 2 IED SIGNAL DETECTION MATRIX

2.2.1.7 Data Requirements.

a.   Number of CO bags screened
b.   Number of MBSs detected (hits) in COs
c.   Configuration of MBSs detected in COs
d.   Number of MBSs not detected (misses) in COs
e.   Configuration of MBSs not detected in COs
f.   Number of MBSs inappropriately suspected (false alarms) in COs
g.   Number of bags passed that do not contain MBSs (correct rejections) in COs
h.   Number of C bags screened
i.   Number of MBSs detected (hits) in C
j.   Configuration of MBSs detected in C
k.   Configuration of MBSs not detected in C
l.   Number of MBSs inappropriately suspected (false alarms) in C
m.   Number of bags passed that do not contain MBSs (correct rejections) in C

2.2.2 Issue 2. Training Effects for Black/White X-Ray Equipment.

Can IED detectability with black/white x-ray equipment be improved with a training intervention?

2.2.2.1 Scope.

Baggage screeners will be trained to detect IEDs using EG&G's off-line TnT computer-based training system. Subsequent baggage screener performance scores in the Post-Training phase will be compared to the Pre-Training phase to determine if the computer-based training improved a screener's probability of detecting an MBS.

2.2.2.2 Criteria.

a. The P_d with CO bags is significantly (statistically) improved after computer-based training.
b. The P_d with C bags is significantly (statistically) improved after computer-based training.
c. The P_{fa} with CO bags is significantly (statistically) reduced.
d. The P_{fa} with C bags is significantly (statistically) reduced.

2.2.2.3 Rationale.

Inferior system performance might reflect a training shortfall rather than equipment inadequacies. Therefore, experience will be given with computer-based IED detection training. This assessment will determine if such training enhances performance in detecting MBSs using black/white x-ray equipment.

2.2.2.4 Evaluation Approach.

The experimental design will allow the assessment of the effectiveness of manufacturer designed computer-based training for enhancing operator capabilities in detecting MBSKs with x-ray based IEDSS. Training effectiveness will be assessed in the operational environment by randomly inserting test bags into the normal flow of screened baggage during the Post-Training phase.

2.2.2.5 MOPs.

MOP 3. The P_d and P_{fa} in C bags for various MBS configurations using black/white x-ray equipment before computer-based training.

MOP 4. The P_d and P_{fa} in CO bags for various MBS configurations using black/white x-ray equipment before computer-based training.

MOP 5. The P_d and P_{fa} in C bags for various MBS configurations using black/white x-ray equipment after computer-based training.

MOP 6. The P_d and P_{fa} in CO bags for various MBS configurations using black/white x-ray equipment after computer-based training.

8

2.2.2.6 Analysis Methodology and Data Presentations.

a. Detection and false alarm rates will be tabulated in a 2 by 2 Signal Detection Matrix and grouped according to C and CO baggage.

b. P_d and P_{fa} will be determined across groups.

c. β and d' values will be calculated for C and CO bags in Pre- and Post-Training phases.

d. Detection rates will be evaluated using the SDT paradigm.

2.2.2.7 Data Requirements.

a. Number of CO bags screened
b. Number of MBSs detected (hits) in COs
c. Configuration of MBSs detected in COs
d. Number of MBSs not detected (misses) in COs
e. Configuration of MBSs not detected in COs
f. Number of MBSs inappropriately suspected (false alarms) in COs
g. Number of bags passed that do not contain MBSs (correct rejections) in COs
h. Number of C bags screened
i. Number of MBSs detected (hits) in C
j. Configuration of MBSs detected in C
k. Number of MBSs not detected (misses) in C
l. Configuration of MBSs not detected in C
m. Number of bags passed that do not contain MBSs (correct rejections) in C

2.2.3 Issue 3. Enhanced X-Ray Detection.

Can baggage screeners better detect MBSs using enhanced x-ray equipment?

2.2.3.1 Scope.

This issue examines the relative effectiveness of IEDSS using enhanced x-ray equipment over those using black/white x-ray equipment to detect MBSs in both C and CO cluttered baggage. Representative passenger bags containing FAA MBS test objects will be inserted into the normal baggage screening operations.

2.2.3.2 Criteria.

a. The P_d with CO bags is significantly (statistically) improved over the P_d with black/white x-ray equipment.

b. The P_d with C bags is significantly (statistically) improved over the P_d with black/white x-ray equipment.

c. The P_{fa} with CO bags is significantly (statistically) reduced.

d.   The P_{fa} with C bags is significantly (statistically) reduced.

2.2.3.3 Rationale.

Enhanced x-rays are used as detection systems in U.S. airports while having an unproved but assumed superiority over black/white x-ray systems. Inserting MBS test objects within cluttered bags into actual baggage flow provides needed information on a system's ability relative to black/white equipment.

2.2.3.4 Evaluation Approach.

The P_d and SDT paradigm measures will be used to assess system performance in the operational environment.

2.2.3.5 MOPs.

MOP 7. The P_d and P_{fa} of various MBS configurations in C bags with enhanced x-ray.

MOP 8. The P_d and P_{fa} of various MBS configurations in CO bags with enhanced x-ray.

2.2.3.6 Analysis Methodology and Data Presentations.

a.   Detection rates will be tabulated in a 2 by 2 signal detection matrix and grouped according to checked and carry-on baggage.

b. P_d and P_{r} . . be determined across groups.

c. β and d' values will be calculated for C and CO bags.

2.2.3.7 Data Requirements.

a.   Number of CO bags screened
b.   Number of MBSs detected (hits) in COs
c.   Configuration of MBSs detected in COs
d.   Number of MBSs not detected (misses) in COs
e.   Configuration of MBSs not detected in COs
f.   Number of MBSs inappropriately suspected (false alarms) in COs
g.   Number of bags passed that do not contain MBSs (correct rejections) in COs
h.   Number of C bags screened
i.   Number of MBSs detected (hits) in C
j.   Configuration of MBSs detected in C
k.   Number of MBSs not detected (misses) in C
l.   Configuration of MBSs not detected in C
m.   Number of MBSs inappropriately suspected (false alarms) in C
n.   Number of bags passed that do not contain MBSs (correct rejections) in C

2.2.4 Issue 4. Training Effects for Enhanced X-Ray Equipment.

Can MBS detectability with enhanced x-ray equipment be improved with a training intervention?

2.2.4.1 Scope.

Baggage screeners will be trained to detect IEDs using EG&G's TnT off-line computer-based training system. Subsequent baggage screener performance scores in the Post-Training phase will be compared to the Pre-Training phase to determine if the computer-based training improved a screener's probability of detecting an MBS.

2.2.4.2 Criteria.

a.   The P_d with C bags is significantly (statistically) improved after computer-based training.

b.   The P_d with CO bags is significantly (statistically) improved after computer-based training.

c.   The P_{fa} with CO bags is significantly (statistically) reduced.

d.   The P_{fa} with C bags is significantly (statistically) reduced.

2.2.4.3 Rationale.

Inferior system performance might reflect a training shortfall rather than equipment inadequacies. Therefore, experience will be given with computer-based IED detection training. This assessment will determine if such training enhances performance in detecting MBSs using enhanced x-ray equipment.

2.2.4.4 Evaluation Approach.

The experimental design will assess the effectiveness of manufacturer-designed computer-based training improving detecting MBSs with enhanced x-ray equipment. Training effectiveness will be assessed in the operational environment by randomly inserting test bags into the normal flow of baggage during the Post-Training phase.

2.2.4.5 MOPs.

MOP 9. The P_d and P_{fa} in C bags for various MBS configurations using enhanced x-ray equipment before computer-based training.

MOP 10. The P_d and P_{fa} in CO bags for various MBS configurations using enhanced x-ray equipment before computer-based training.

MOP 11. The P_d and P_{fa} in C bags for various MBS configurations using enhanced x-ray equipment after computer-based training.

MOP 12. The P_d and P_{fa} in CO bags for various MBS configurations using enhanced x-ray equipment after computer-based training.

2.2.4.6 Analysis Methodology and Data Presentations.

a. Detection rates will be tabulated in a 2 by 2 signal detection matrix and grouped according to C and CO baggage both prior and following training.

b. P_d and P_{fa} will be determined across groups.

c. β and d' values will be calculated for C and CO bags in the Pre- and Post-Training phases.

d. Detection rates will be evaluated using the SDT paradigm.

2.2.4.7 Data Requirements.

a. Number of CO bags screened
b. Number of MBSs detected (hits) in COs
c. Configuration of MBSs detected in COs
d. Number of MBSs not detected (misses) in COs
e. Configuration of MBSs not detected in COs
f. Number of MBSs inappropriately suspected (false alarms) in COs
g. Number of bags passed that do not contain MBSs (correct rejections) in COs
h. Number of C bags screened
i. Number of MBSs detected (hits) in C
j. Configuration of MBSs detected in C
k. Number of MBSs not detected (misses) in C
l. Configuration of MBSs not detected in C
m. Number of MBSs inappropriately suspected (false alarms) in C
n. Number of bags passed that do not contain MBSs (correct rejections) in C

2.3   DATA SOURCE MATRIX.

Table 3 shows the primary (P) sources for all data elements to be collected.

TABLE 3. PRIMARY (P) DATA SOURCES

Data Element	IEDDS OT&E	Other
Number of bags screened	P	
Number of MBSs detected (hits)	P	
Number of MBSs not detected (misses)	P	
Number of MBSs inappropriately suspected (false alarms)	P	
Number of bags passed as not containing MBSs (correct rejections)	P	
P_d	P	
P_d for MBS configurations in CO uncluttered bags		P

2.3.1   Issue 1.

Detection data to calculate P_d will be collected to determine what extent x-ray screeners can detect MBSs in cluttered bags using black/white x-ray equipment for both C and CO baggage. Signal detection elements (β and d') will also be calculated to determine baggage screener MBS detection characteristics using black/white x-ray equipment. The C phase data will be contrasted with that recently collected by the FAA using the same MBS configurations in a non-cluttered C bag.

2.3.2   Issue 2.

Data will be collected to determine if MBS detection rates with black/white x-ray equipment can be significantly improved with training. The Pre-Training and Post-Training P_d and P_{fa} will be analyzed within groups.

2.3.3 Issue 3.

Detection data to calculate P_d and P_{fa} will be collected to determine if x-ray screeners can detect MBSs in cluttered bags significantly better using enhanced x-ray equipment rather than black/white x-ray equipment. Signal detection elements (β and d') will also be calculated to determine baggage screener MBS detection characteristics using enhanced x-ray equipment. The CO phase data will be contrasted with that recently collected by the FAA using the same MBS configurations in a non-cluttered bag.

2.3.4 Issue 4.

Data will be collected to determine if detection rates with enhanced x-ray equipment can be improved with training. The Pre-Training and Post-Training P_d will be analyzed within groups.

2.4 TEST APPROACH.

2.4.1 Test Scope.

The scope of this test includes those activities necessary to determine MBS detection and false alarm rates with black/white and enhanced x-ray equipment, for C and CO baggage and various MBS configurations. The test also includes an assessment of IEDSS performance following computer-based training.

2.4.2 Factors and Conditions.

Table 4 lists the IEDSS OT&E factors, control procedures, and conditions.

TABLE 4. OT&E FACTORS, CONTROL PROCEDURES, AND CONDITIONS

Factors	Control	Conditions
Training Level	Systematically Varied	Pre-Training, Post-Training
IEDSS Configuration	Systematically Varied	Black/white or enhanced images
Screeners	Grouped	Years Experience Gender Age Education Level Position
Baggage Comparison:		
Season	Held Constant	Summer
Itinerary	Randomly Varied	Short, Medium, Long
Profile	Randomly Varied	Business, Tourist, Etc.
Airline	Held Constant	United Airlines (UA)
Airport	Held Constant	San Francisco, CA (SFO)
Security Company	Held Constant	ITS
MBS placement	Randomly Varied	20 among 200 C comparison bags or 12 among 120 CO comparison bags

2.4.3 Sample Size.

There will be 40 screener participants selected from those available at SFO. They will be assigned to one of four groups, as previously outlined in section 2.1.1 (Evaluation Approach).

2.4.4 Additional Considerations.

a. The outcomes of statistical tests will be evaluated at an alpha level of 0.05. To correct for the positive bias of the F-test associated with within-subjects effects, the Geisser-Greenhouse ε correction will be used.

b. The TER will be divided into four primary sections. The first will describe the performance of screeners in detecting MBSs on both C and CO conditions for Pre-Training baseline phase. The second section will describe screener performances on both conditions for the Post-Training phase. The third section will compare screener performances for Pre-Training and Post-Training. The difference in scores from the Pre-Training to the Post-Training phases will identify the

training context effect (i.e., carry-over) from the computer-based training. A fourth section will compare the CO OT&E results to data recently collected by the FAA using the same MBS configurations in non-cluttered CO bags.

c.   An outline of the TER is included in appendix A.

2.5   EVALUATION DATABASE STRUCTURE.

The final detection database will consist of the performance scores on the Pre-Training and Post-Training for C and CO bags examined with black/white and enhanced x-ray. The primary statistical software tool that will be used to analyze group performances is Statistica for Windows, Release 4.0. Integrated into the detection database will be appropriate data for trial, personnel, β, d', MBS detection rates, and MBS baggage configurations. The data described in table 5 will be collected and integrated into the signal detection data file.

TABLE 5. DATABASE STRUCTURE

Screener ID
Screener Characteristics
Date
C or CO
Black/white or Enhanced
Training Level
MBS Configuration
Number of Hits
Number of Misses
Number of False Alarms
Number of Correct Rejections
d'
β

3. TEST DESIGN.

3.1 TEST CONCEPT.

3.1.1 Introduction.

The purpose of this OT&E is to assess the operational effectiveness of both standard (black/white) and enhanced ("color") display x-ray based IEDSS in allowing baggage screeners to detect a MBS in a cluttered bag. The OT&E will also focus on whether operational effectiveness is enhanced with computer-based IED detection training. Screeners will be assessed at Pre-Training and Post-Training levels using the P_d, P_{fa}, and a SDT paradigm. Further details on the SDT paradigm can be found in appendix B.

3.1.2 Operational Context.

The OT&E will occur in July and August 1994, at UA's international C and domestic CO baggage locations in SFO. To complete the test stages (i.e., Familiarization, Pilot Test, and Operational Test), the study will require approximately 2-3 weeks (see section 3.1.3 for explanation of testing stages).

3.1.2.1 Scenario.

a. Twenty C baggage screeners will scan 440 bags (i.e., 220 Pre-Training and 220 Post-Training) using black/white (10 screeners) or enhanced (10 screeners) x-ray based IEDSS in their operational setting. The same twenty C MBS test kits will be presented at both the Pre- and Post-Training level. A single x-ray screening station will be used because only one station (number 2 belt) allows for ready retrieval of test bags. These bags will be uniquely numbered and plainly labeled for data collection and retrieval. Barriers will be placed on both ends of the E-Scan x-ray equipment to prevent screeners from seeing the bag before and immediately after the screening process.

b. Twenty CO baggage screeners will scan 264 bags (i.e., 132 Pre-Training and 132 Post-Training) in their operational setting. The same twelve CO MBS test kits will be presented during both the Pre- and Post-Training phases. The single E-Scan machine in domestic check-in will be used for this portion of the test.

c. MBS test kits (MBSKs), for both black/white and enhanced x-ray of C bags, will be randomly inserted into the passenger baggage flow by contractor personnel. The CO MBSKs will be carried by FAA personnel randomly inserted into passenger flow.

d. Baggage screeners will be informed that threat IED test objects may be presented. The screeners' MBS detection data will be recorded as a hit (correctly detects MBSK), miss (fails to detect MBSK and passes bag through), false alarm (incorrectly suspects a MBSK when none is present), or correct rejection (passes a bag through when no MBSK is present). Figure 4 displays the flow of the OT&E operations for each screener response.

e. FAA personnel will stand next to each screener when test bags are being inserted. They will immediately determine whether any given bag, thought to potentially contain an IED, by a

17

screener, is a FAA test bag. This decision for C bags will be based on inspecting for a fluorescent yellow external marker. The CO test bags will be identified by examining bags for ultra-violet markings responsive to black light.

f.   The OT&E will also determine if computer-based training will enhance the IEDSS' performance in detecting MBSKs. The training will be on the EG&G TnT system which the FAA Technical Center will ship before the operational test.
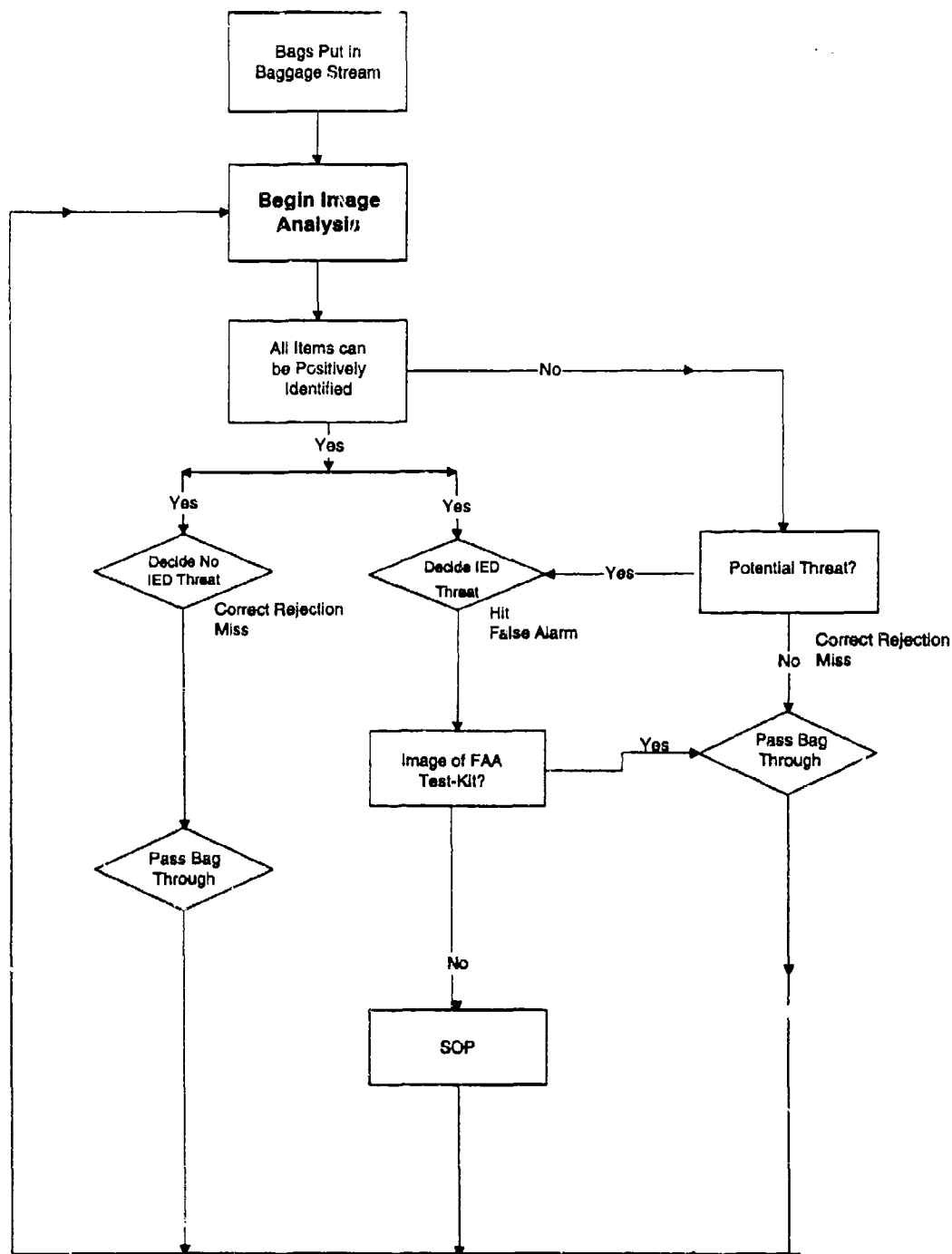
FIGURE 3. THE FLOW OF OT&E OPERATIONS FOR EACH SCREENER RESPONSE

19

3.1.3 Test Stages.

The OT&E will be conducted in three stages: Familiarization, Pilot Testing, and Operational Testing.

3.1.3.1 Familiarization.

Before the OT&E, EG&G Astrophysics will train FAA and contractor personnel in the operation of the TnT training system, at the FAA Technical Center Aviation Security Laboratory.

3.1.3.2 Pilot Testing.

Before the OT&E, FAA and contractor personnel will conduct pilot tests to ensure that data collection procedures are correct and to further coordinate logistical issues.

a. The pilot-I test (15 June 1994) will use several C and CO bags, without MBSKs, to examine logistical procedures. Issues will include identification of test bags for C and CO conditions, retrieveability of C bags from the conveyer belt, and viability of back-up retrieval system relying on baggage tags.

b. The pilot-II test (25-26 July 1994) will include screeners and MBSKs. Two baggage screeners will participate in the study. Screeners will view the same C and CO MBSKs to be used in the operational test.

3.1.3.3 Operational Testing.

a. The operational test will occur immediately after pilot-II testing and will take approximately 2-3 weeks.

b. FAA MBSKs will be inserted into routine C and CO baggage flow. The MBSKs for CO bags will have three configurations previously found to have different P_ds (i.e., low, medium, and high) in uncluttered bags. Ten MBSK configurations will be used for C bags. Section 3.1.7 describes these MBS configurations.

3.1.4 Test Unit Configuration.

a. Test personnel will include FAA, contractor employees (GSC and DCS), as well as ITS security screeners.

b. Baggage screeners will scan bags with an EG&G Astrophysics E-scan system with enhancements turned off for the black/white x-ray equipment evaluation.

c. Screeners will scan bags with an EG&G Astrophysics E-scan System with black and white as well as enhancements available for the enhanced x-ray equipment evaluation.

d. System performance on both black/white and enhanced x-ray systems will be tested during the Pre- and Post-Training phases.

e.   Individual screener performance will not be reported.

3.1.5 Training Concept.

Appendix C describes the TnT operating procedures.

3.1.5.1 Briefing Direct and Indirect Participants.

Before data collection, FAA and contractor personnel will brief screeners and essential airline, airport, security company, and law enforcement personnel about the OT&E's purpose, schedule, and procedures. The briefings will occur during the preliminary coordination trip (24 May 1994) to SFO and as part of the Operational Test Readiness Review (OTRR) (16 June 1994).

3.1.5.2 Training Test Subjects.

The test subjects (x-ray screeners) will receive operational instructions and computer-based training using appropriate modules of the TnT system. The selected modules will address detecting IED threats only.

3.1.5.3 Training the Test Organization.

Training for FAA and contractor test team personnel will occur before the pilot-II test. The training session will take 1 day and cover data collection procedures and methodology.

3.1.6 Overall Methodology.

a.   A mixed-factorial pre-post design will be used to assess system performance before and after a computer-based training intervention.

b.   Forty baggage screeners (i.e., four groups of baggage screeners with 10 screeners per group) will use either black/w ite or enhanced x-ray equipment for C or CO bags.

c.   Each C bag screener will see approximately 440 images (i.e., 220 images per phase). The images will come from 400 real passenger bags and 40 FAA MBSKs (i.e., 20 MBSKs per phase).

d.   Each CO bag screener will see approximately 264 images (i.e., 132 images per phase). The images will come from 240 real passenger bags and 24 FAA MBSKs (i.e., 12 MBSKs per phase).

e.   These signal presentation rates of 10 percent are presumably much higher than actual threat percentages. However, d' is independent of this.

f.   Baggage screeners will attempt to detect the presence of a MBSK. Since there are only two possible states (i.e., the MBSK was present or not present), there are four possible response outcomes: Hits, Misses, False Alarms, and Correct Rejections. Actual threat objects, if any, will not be included in these calculations.

g.   The FAA will ship the stand-alone TnT system to SFO before conducting pilot-II and operational testing. Pilot testing will ensure that all equipment functions properly: that test participants understand their role, and that test team participants are able to successfully obtain the required data.

h.   The groups of test subjects (C-black/white, C-enhanced; CO-black/white, CO-enhanced) will take computer-based training on the TnT system after the Pre-Training phase. Subjects will work individually on the TnT to complete the required modules. All screeners will also be given a test of visual contrast sensitivity. Subjects in the enhanced x-ray condition will additionally receive a color vision test. The TnT training and vision testing time will take approximately 1-1.5 hours.

i.   The support contractor will pay screeners for participating in the training session.

j.   Demographic data will be collected and related to performance.

3.1.7   Test Bag Composition.

3.1.7.1 Representativeness.

a.   Representativeness of C and CO bags used in the OT&E will be based upon findings from a previous FAA study. That study assessed approximately 3,507 C passenger bags, at Miami International Airport, in terms of such variables as clutter, density, and organic compound content. The OT&E will use small- (less than 22 inches), medium- (22 - 28 inches), and large-sized (28–32 inches) bags.

b.   Table 6 presents means and standard deviations for the variables used to select representative bags to contain MBSXs for each baggage size category. Test bags will be selected for each size category that have the same (+/- 1 SD) amount of clutter, density, and organic content as found in the average bag for the Miami analysis.

c.   The C bags will include the same percentages of small (11 percent), medium (47 percent), and large-sized (42 percent) bags found in the Miami study.

d.   Only small sized bags (less than 22 inches) will be as CO bags.

TABLE 6. MEAN AND STANDARD DEVIATIONS OF THE MIAMI INTERNATIONAL
CHECKED BAGS AS A FUNCTION OF CLUTTER, DENSITY AND ORGANIC CONTENT

Size	Clutter	Density	Organic
Small < 22 " n (C) = 2 n (CO) = 12	$\overline{X}$ = 5.4 SD = 4.2	$\overline{X}$ = 5.9 SD = 2.6	$\overline{X}$ = 11,119 SD = 6,289
Medium 22" - 28" n (C) = 10	$\overline{X}$ = 3.4 SD = 2.4	$\overline{X}$ = 5.3 SD = 1.6	$\overline{X}$ = 16,245 SD = 6,432
Large 28" - 32" n (C) = 8	$\overline{X}$ = 3.3 SD = 2.0	$\overline{X}$ = 5.5 SD = 1.3	$\overline{X}$ = 24,095 SD = 7,090

3.1.7.2 MBSK Configuration.

a.   The CO test bags will contain MBSKs similar to three of the ten MBS configurations used by the FAA (ACO-130) in a previous assessment of MBS detection using uncluttered CO bags. The three configurations to be used correspond to the highest, lowest, and intermediate P_d observed in that assessment. The configurations are as follows: C-4, Polaroid battery, and digital travel alarm; TnT, 9-volt battery, and analog travel clock; and C-4, 21.5 battery, and analog travel clock.

b.   The C bags will contain MBSKs in each of the same 10 configurations used by the FAA (ACA-130) in a previous assessment of MBS detection using uncluttered CO bags. These configurations include explosive simulants (dynamite, TNT, C-4, sheet), batteries (Polaroid, 9-volt, two 1.5-volt), and timers (small, medium, and large analog, digital).

c.   The C gags will include the same percentages of small- (11 percent), medium- (47 percent), and large-sized (42 percent) bags found in the Miami study.

d.   Only small-sized bags (less than 22 inches) will be used as CO bags.

e.   Blank detonators will be included in all simulated explosives for all configurations.

3.1.8   Test Equipment and Materials.

Instrumentation required for the OT&E includes the black/white and enhanced x-ray equipment (EG&G), FAA MBSKs, EG&G TnT trainer, black light source, counters, flashlight, portable computer and analytical software, and data collection forms. The x-ray equipment will be the EG&G equipment currently used at SFO.

3.1.9   Test Limitations.

a.   Baggage screeners will be selected based upon their availability which does not ensure a representative sample.

b.   The black/white x-ray condition consists of an EG&G x-ray machine with enhancements off. This may not adequately reflect the capability of all black/white machines.

c.   Both the fact that the operators will be aware their performance is under inspection and threat items will occur at a high rate can effect performance. Conditions which alter arousal level and/or attention to the task effect β. Conclusions drawn from data collection in this fashion may be affected by an artificially high β.

3.2   TEST DETAILS.

3.2.1   Issue 1. Black/White X-Ray Detection.

Can baggage screeners detect MBSs using black/white x-ray equipment?

3.2.1.1 Operational Test Measure of Performance (OTMOP 1).

The P_d and P_{fa} of each MBS configuration in C bags with black/white x-ray.

3.2.1.1.1 Data Collection Methodology.

Data collection methodology includes the following:

a.   Twenty uniquely numbered bags containing MBSKs will be used as C baggage test items. These test bags will contain simulated IEDs in one of ten different configurations and will be randomly distributed among 200 passenger C bags.

b.   Ten baggage screeners will scan C bags using black/white x-ray equipment.

c.   Contractor personnel will randomly insert test bags among passenger bags and recover C test bags out of sight of the screener. FAA personnel will be stationed next to each baggage screener

24

and record the screener's decision on whether the bags may contain a MBSK. For suspect bags, the screener will immediately be told whether it is a FAA test bag. Test bags will include unique bag numbers externally marked on fluorescent yellow stickers. A third person will collect MBSKs from the number two conveyor.

d.   Standard security procedures will be followed for suspicious passenger bags that are not FAA test bags.

e.   Bar codes will also be used in C MBSKs to provide redundant identification. A unique number on the bar code assures positive bag identification if external bag numbers are unidentifiable.

f.   An internal written statement will additionally identify C test bags as belonging to the FAA test.

3.2.1.1.2 Data Requirements.

Data requirements include the following:

a.   A 2 by 2 SDT matrix (See figure 2) of baggage screener responses for each MBS configuration used.

b.   Totals of each cell and both columns in the 2 by 2 matrix.

c.   The P_d and P_{fa} for MBSKs in C bags (see appendix B for calculations).

3.2.1.1.3 Data Reduction and Analyses.

Data Reduction and Analyses include the following:

a.   The 2 by 2 matrix will be used to calculate all probabilities. The P_d will be calculated by dividing the number of MBSs detected by the total number of MBSKs. See appendix B for further explanation on calculating SDT elements.

b.   The P_{fa} will be determined by dividing the number of false alarms by the total number of innocent bags.

c.   The probability of missing a MBS (P_m) and the probability to correctly pass a bag that does not contain an MBS (i.e., correct rejection, P_{cr}) will be calculated by the following equations: $P_m = 1 - P_h$; $P_{cr} = 1 - P_{fa}$.

3.2.1.2 OTMOP 2.

The P_d and P_{fa} of MBS configurations for CO bags with black/white x-ray.

3.2.1.2.1 Data Collection Methodology.

Data collection methodology includes the following:

a.   Twelve bags containing a MBSK will be used as CO baggage test items. These bags will contain simulated IEDs, in one of three configurations (see section 3.1.7), randomly distributed among 120 passenger CO bags.

b.   Ten baggage screeners will scan CO bags using black/white x-ray equipment.

c.   FAA personnel will bring the CO bags, through the single CO screening station with an EG&G enhanced x-ray, for black/white screening. After passing through the screening station, test bag carriers will return the bag to the staging area. FAA personnel will be stationed next to each baggage screener and record the screener's decision on whether the bags may contain a MBSK. Suspect bags will be examined for ultra-violet markings sensitive to black light and the screener will immediately be told whether it is a FAA test bag.

d.   Standard security procedures will be followed for suspicious passenger bags that are not FAA test bags.

e.   An internal written statement will additionally identify CO test bags as belonging to the FAA test.

3.2.1.2.2 Data Requirements.

Data requirements include the following:

a.   A 2 by 2 SDT matrix of baggage screener response for each of the three MBSK configurations.

b.   Totals of each cell and both columns in the 2 by 2 matrix.

c.   The P_d and P_{fa} for MBSKs in CO bags.

3.2.1.2.3 Data Reduction and Analyses.

Data Reduction and Analyses include the following:

a.   The 2 by 2 matrix will be used to calculate all probabilities. The P_d for Pre-Training scores will be calculated by dividing the number of MBSs detected by the total number of MBSKs detected and missed. See appendix B for further explanation on calculating SDT elements.

b.   The P_{fa} will be determined by dividing the number of false alarms by the total number of false alarms and correct rejections .

c.   The P_m and P_{cr} will be calculated by the following equations: $P_m = 1 - P_h$; $P_{cr} = 1 - P_{fa}$.

3.2.1.3 OTMOP 3.

β and d' for C bags with black/white x-ray.

3.2.1.3.1 Data Collection Methodology.

Data collection methodology is the same as that for OTMOP 1.

3.2.1.3.2 Data Requirements.

Data Requirements include the following:

a. A 2 by 2 SDT matrix of baggage screener responses for each configuration.

b. Totals of each cell and both columns in the 2 by 2 matrix.

c. Combined screener's β and d' (See appendix B for explanation of calculations).

3.2.1.3.3 Data Reduction and Analyses.

Data reduction and analyses includes data recorded in 2 by 2 SDT matrix and combined for all screeners. β and d' will be calculated for all screeners at each MBS configuration.

3.2.1.4 OTMOP 4.

β, and d' for CO bags with black/white x-ray.

3.2.1.4.1 Data Collection Methodology.

Data collection methodology is the same as that for OTMOP 3.

3.2.1.4.2 Data Requirements.

Data Requirements are the same as those for OTMOP 3.

3.2.1.4.3 Data Reduction and Analyses.

β and d' as in OTMOP 3.

3.2.2 Issue 2. Training Effects for Black/White X-Ray Equipment.

Can MBS detectability with black/white x-ray equipment be improved with a training intervention?

3.2.2.1 OTMOP 5.

The P_d and P_{fa} values for MBSs in C bags, using black/white x-ray equipment before computer-based training, as calculated in OTMOP 1.

3.2.2.2 OTMOP 6.

The P_d and P_{fa} values for MBSs in CO bags, using black/white x-ray equipment before computer-based training, as calculated in OTMOP 2.

3.2.2.3 OTMOP 7.

The P_d and P_{fa} of MBSs in C bags using black/white x-ray equipment after computer-based training. Procedures are the same as those for OTMOP 1.

3.2.2.4 OTMOP 8.

The P_d and P_{fa} of MBSs in CO bags using black/white x-ray equipment after computer-based training. Procedures are the same as those for OTMOP 2.

3.2.2.5 OTMOP 9.

β and d' values for C bags, before training, as calculated in OTMOP 3.

3.2.2.6 OTMOP 10.

β and d' values for C bags after training. Procedures are the same as those for OTMOP 3.

3.2.2.7 OTMOP 11.

β and d' values, for CO bags before training, as calculated in OTMOP 4.

3.2.2.8 OTMOP 12.

β and d' values for CO bags after training. Procedures are the same as those for OTMOP 4.

3.2.3 Issue 3. Enhanced X-Ray Detection.

Can baggage screeners better detect MBSs using enhanced x-ray equipment?

3.2.3.1 OTMOP 13.

The P_d and P_{fa} of MBS configurations in C bags.

3.2.3.1.1 Data Collection Methodology.

Data collection methodology includes the following:

a. Twenty uniquely numbered bags containing MBSKs will be used as C baggage test items. These bags will contain simulated IEDs in one of ten different configurations and will be randomly distributed among 200 passenger C bags.

b. Ten baggage screeners will scan C bags using enhanced x-ray equipment.

c. Contractor personnel will randomly insert test bags among passenger bags and recover C test bags. One person will randomly place test-bags on the conveyor belt out of sight of the screener. FAA personnel will be stationed next to each baggage screener and record the screener's decision on whether the bags may contain a MBSK. For suspect bags, the screener will immediately be told whether it is a FAA test bag. Test bags will include unique bag numbers externally marked on fluorescent yellow stickers. A third person will collect MBSKs from the number two conveyor.

d. Standard security procedures will be followed for suspicious passenger bags that are not FAA test bags.

e. Bar codes will also be used in C MBSKs to provide redundant identification. A unique number on the bar code assures positive bag identification if external bag numbers are unidentifiable.

f. An internal written statement will additionally identify C test bags as belonging to the FAA test.

3.2.3.1.2 Data Requirements.

Data requirements include the following:

a. A 2 by 2 SDT matrix of baggage screener responses for each MBS configuration used.

b. Totals of each cell and both columns in the 2 by 2 matrix.

c. The P_d and P_{fa} for MBSKs.

3.2.3.1.3 Data Reduction and Analyses.

Data Reduction and Analyses include the following:

a. The 2 by 2 matrix will be used to calculate all probabilities. The P_d will be calculated by dividing the number of MBSs detected by the total number of MBSKs. See appendix B for further explanation on calculating SDT elements.

b. The P_{fa} will be determined by dividing the number of false alarms by the total number of innocent bags.

c. The probability of missing a MBS (P_m) and the probability to correctly pass a bag that does not contain an MBS (i.e., correct rejection, P_{cr}) will be calculated by the following equations: $P_m = 1 - P_h$; $P_{cr} = 1 - P_{fa}$.

3.2.3.2 OTMOP 14.

T e P_d and P_{fa} of MBS configurations for CO bags.

3.2.3.2.1 Data Collection Methodology.

Data collection methodology includes the following:

a. Twelve bags containing a MBSK will be used as CO baggage test items. These bags will contain simulated IEDs, in one of three configurations (see section 3.1.7), and will be randomly distributed among 120 passenger CO bags.

b. Ten baggage screeners will scan CO bags using enhanced x-ray equipment.

c. FAA personnel will bring the CO bags, through the single CO screening station with an EG&G enhanced x-ray, for enhanced screening. After passing through the screening station, test bag carriers will return the bag to the staging area. FAA personnel will be stationed next to each baggage screener and record the screener's decision on whether the bags may contain a MBSK. Suspect bags will be examined for ultra-violet markings sensitive to black light and the screener will immediately be told whether it is a FAA test bag.

d. Standard security procedures will be followed for suspicious passenger bags that are not FAA test bags.

e. An internal written statement will additionally identify CO test bags as belonging to the FAA test.

3.2.3.2.2 Data Requirements.

Data requirements include the following:

a. A 2 by 2 SDT matrix of baggage screener response for each of the three MBSK configurations.

b. Totals of each cell and both columns in the 2 by 2 matrix.

c. The P_d and P_{fa} for each MBSK configuration.

3.2.3.2.3 Data Reduction and Analyses.

Data Reduction and Analyses include the following:

a. The 2 by 2 matrix will be used to calculate all probabilities. The P_d for Pre-Training scores will be calculated by dividing the number of MBSs detected by the total number of MBSKs detected and missed. See appendix B for further explanation on calculating SDT elements.

b. The P_{fa} will be determined by dividing the number of false alarms by the total number of false alarms and correct rejections .

c. The P_m and P_{cr} will be calculated by the following equations: $P_m = 1 - P_h$; $P_{cr} = 1 - P_{fa}$.

3.2.3.3 OTMOP 15.

β, and d' values for C bags. Procedures are the same as those for OTMOP 3.

3.2.3.4 OTMOP 16.

β, and d' values for CO bags. Procedures are the same as those for OTMOP 4.

3.2.4 Issue 4 Training Effects for Enhanced X-Ray Equipment.

Can MBS detectability with enhanced x-ray equipment be improved with a training intervention?

3.2.4.1 OTMOP 17.

The P_d and P_{fa} values for MBSs in C bags, using enhanced x-ray equipment before computer-based training, as calculated in OTMOP 13.

3.2.4.2 OTMOP 18.

The P_d and P_{fa} values for MBSs in CO bags, using enhanced x-ray equipment before computer-based training, as calculated in OTMOP 14.

3.2.4.3 OTMOP 19.

The P_d and P_{fa} for MBSs in C bags using enhanced x-ray equipment after computer-based training. Procedures are the same as those for OTMOP 17.

3.2.4.4 OTMOP 20.

The P_d and P_{fa} for MBSs in CO bags using enhanced x-ray equipment after computer-based training. Procedures are the same as those for OTMOP 18.

3.2.4.5 OTMOP 21.

β, and d' values, for C bags before training, as calculated in OTMOP 15.

31

3.2.4.6 OTMOP 22.

β, and d' values for C bags after training. Procedures are the same as those for OTMOP 15.

3.2.4.7 OTMOP 23.

β, and d' values, for CO bags before training, as calculated in OTMOP 16.

3.2.4.8 OTMOP 24.

β, and d' values for CO bags after training. Procedures are the same as those for OTMOP 16.

4. REFERENCES.

1. Coren, S. and L.M. Ward, *Sensation and Perception*, 3rd Edition, New York, Harcourt Brace Jovanovich, 1989.

2. Wickens, C.D., *Engineering Psychology and Human Performance*, 2nd Edition, New York, Harper-Collins, Inc., 1992.

DOT/FAA/CT-xx/xx

Appendix A–Outline Test and Evaluation Report (TER) for the Improvised Explosive Device Screening Systems (IEDSS)

FAA Technical Center
Atlantic City Int'l Airport,
N.J. 08405

August 1994

Draft

U. S. Department of Transportation

Federal Aviation Administration

INTRODUCTION
    Purpose
    Scope
    Background
        Program Background
        Test and Evaluation Background
    System Description
    Operational Issues and Criteria
        COIC 1
        COIC n
    Evaluation Approach
    Test and Evaluation Limitations and Impact
        Test Limitations and Impact
        Evaluation Limitations and Impact
TEST DESCRIPTION
    Test Purpose
    Test Overview
        System Familiarization
        Pilot Test
        Operational Test
    Conduct of Test
        Test events
        Control Procedures
        Schedule of Events
        Overall Methodology
    Test Data Management

        Description of Data Collection Methods

        Database

    Training Conducted

        Security Company Employees

TEST AND EVALUATION RESULTS

    Issue 1:

        Criterion 1.1

        Criterion 1.n

        Issue Evaluation and Conclusions

    Issue n:

        Criterion 2.n

        Issue Evaluation and Conclusions

OVERALL CONCLUSIONS AND RECOMMENDATIONS

    Overall Conclusions and Recommendations

COIC 1

COIC n

Operational Suitability Conclusions

COIC n

Recommendations

System Design Improvements

Future Test and Evaluation

Fielding Concerns

APPENDICES

LIST OF ILLUSTRATIONS

LIST OF TABLES

APPENDIX B

SIGNAL DETECTION THEORY (SDT) AND APPLICATION

SIGNAL DETECTION THEORY AND APPLICATION

The Signal Detection Theory Paradigm.

The IEDDS operation features human operators engaged in tasks to detect an environmental event or signal. Signal Detection Theory (SDT) is a mathematical representation of human performance in deciding whether or not a signal is present. An operational example of SDT is an airport security guard screening passenger bags for IEDs.

There are two response categories that represent a screener's detection performance: Yes (a MBS signal was present) or No (a MBS signal was not present). There are also two signal presentation states indicating that the MBS signal was present (signal) or absent (noise). A combination of security guard responses and the signal state produces a 2 by 2 matrix (figure B-1), generating four classes of operator responses, labeled hits, misses, false alarms, and correct rejections (Wickens). Considering the IEDSS OT&E:

a.   A Hit will be recorded when a baggage screener detects a MBS in the scanned baggage.

b.   A Miss will be recorded when a baggage screener fails to detect a MBS which is present in the scanned baggage and passes the bag through the x-ray equipment.

c.   A False Alarm will be recorded when a baggage screener detects a MBS in the scanned baggage when none is present.

d.   A Correct Rejection will be recorded when a baggage screener passes a bag through that does not contain a MBS.

State of MBS Image

MBS Present   MBS Not Present

Test Subject
Response

Yes   Hit   False Alarm

No   Miss   Correct
                                 Rejection

FIGURE B-1. 2 BY 2 MATRIX OF SECURITY GUARD RESPONSES
AND STATE OF MBS IMAGE

As indicated by Wickens, the SDT paradigm assumes that operators perform two stages of information processing in all detection tasks: (1) sensory evidence is aggregated concerning the presence or absence of the signal, and (2) a decision is made about whether this evidence constitutes a signal or not. According to SDT, external stimuli generate neural activity in the brain. On the average, there will be more sensory or neural evidence in the brain when a signal is present than when it is absent. This neural evidence, X, referred to as the evidence variable, represents the rate of firing of neurons in the brain. The response rate for detecting X increases in magnitude with stimulus (signal) intensity. Therefore, if there is enough neural activity, X exceeds a critical threshold, X_c, and the operator decides "yes." If there is too little, the operator decides "no." Because the amount of energy in the signal is typically low, the average amount of X generated by signals in the environment is not much greater than the average generated when no signals are present (noise). Furthermore, the quantity of X varies continuously, even in the absence of a signal, because of random variations in the environment and the operator's level of neural firing (i.e., the neural "noise" in the operator's sensory channels and brain).

The relationship between the presence and absence of a signal, random variability of X, and X_c can be seen in hypothetical noise and signal distributions (figure B-2). Figure B-2 plots the probabilities of

B-3

observing a specific value of X, given that a noise trial (left distribution) or signal trial (right distribution) occurred. The intersection of the two curves represents the location where the probability of a signal equals the probability of noise. The criterion value, X_c, chosen by the operator is shown by the vertical line. All X values to the right ($X > X_c$) will cause the operator to respond "yes." All X values to the left generate "no" responses. The different shaded areas represent the occurrences of hits, misses, false alarms, and correct rejections.
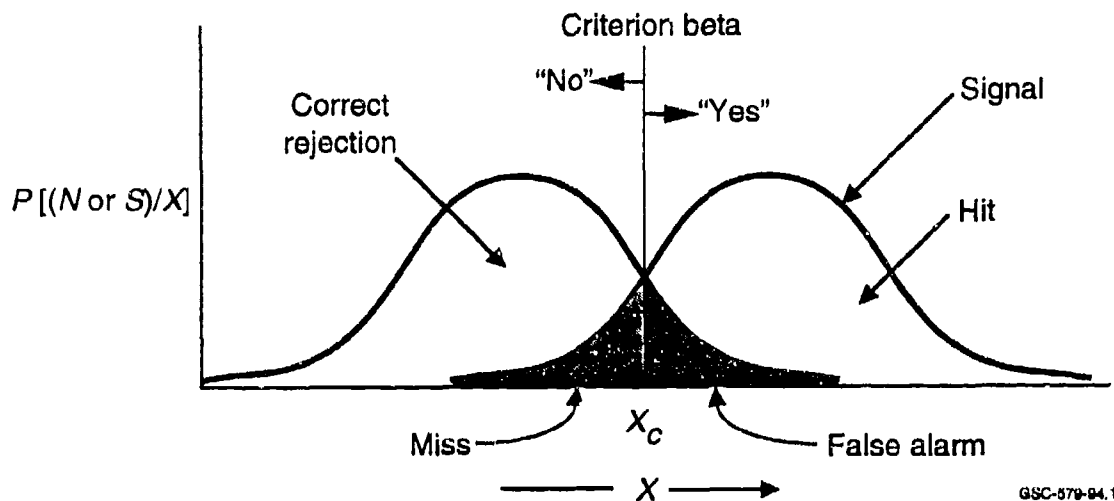

FIGURE B-2. HYPOTHETICAL SDT DISTRIBUTIONS (Wickens, 1992)

Procedures to Calculate SDT Probabilities.

a.   In SDT, the detection values are expressed as probabilities;

b.   The probability of hit (P_h), miss (P_m), false alarm (P_{fa}), and correct rejection (P_{cr}) are determined by dividing the number of occurrences in a cell (figure B-1) by the total number of occurrences in a column;

c.   The P_h (also referred to as probability of detection (P_d)) will be calculated by dividing the number of IEDs detected (number of hits) by the total number of hits and misses;

d.   The P_{fa} will be determined by the number of false alarms divided by the total number of false alarms and correct rejections;

e.   Since the total area within each curve equals one, the sum of the two shaded regions within each curve must also equal one. That is, $P_h + P_m = 1$ and $P_{fa} + P_{cr} = 1$.

B-4

<u>Operator Response Criterion</u>.

In any signal detection task, operator decision making may be described in terms of an operator response criterion. Operators may use "risky" response strategies by responding yes more often than no. A risky strategy allows operators to detect most of the signals that occur, but also produces many false alarms. Alternatively, operators may use "conservative" strategies, saying no most of the time, making few false alarms, but missing many of the signals.

Different circumstances may require conservative or risky strategies. For example, an appropriate IED detection strategy requires screeners to respond yes when there is question of baggage contents. This response may produce false alarms when no threatening objects are present.

As shown in figure B-2, risky or conservative behavior is determined by the location of the operator's response criterion, X_c. If X_c is placed to the right, much evidence of the signal is required for it to be exceeded and most responses will be no (conservative responding). If it is placed to the left, little signal evidence is required and most responses will be yes, or risky. A variable positively correlated with X_c is the quantity beta (β), which is the ratio measure of operator neural activity utilized to produce a response:

$$\beta = \frac{P(X/S)}{P(X/N)}$$

This equation is the ratio of the ordinate of the two curves of figure B-2, at a given level of X_c. The higher β values will generate fewer yes responses and, therefore, fewer hits. Lower β settings will generate more yes responses, more hits, and more false alarms.

The actual probability values appearing in the 2 by 2 matrix (figure B-1) determine the value of β. The probabilities define the areas under the two distribution functions shown in figure B-2, to the left and right of the criterion. Thus, the P_d is the relative area under the signal curve (a signal was present) to the right of the criterion (the operator said yes).

Table B-1 provides a representative table of Z values and ordinate values of the probability distribution related to hit and false alarm responses. A complete table of the area under the standard normal distribution will be used to calculate β for the Test and Evaluation Report (TER). The procedures required to calculate β are listed below (Coren and Ward).

<u>Procedures to Calculate β:</u>

a. Find the false alarm rate from the outcome matrix in the HIT/FA column of table B-1;
b. Read across the table to the ORD column (for ordinate, the height of the bell curve);
c. Calculate the value tabled there ORD(FA) and write it down;
1. Repeat these operations for the hit rate, calling the tabled value ORD(HIT), and write it down;
e. Calculate β using the following equation: $\beta = ORD(HIT)/ORD(FA)$;

Sensitivity (d').

Sensitivity refers to the average amount of operator sensory activity generated by a given signal as compared with the average amount of noise-generated activity (Coren and Ward). As explained earlier, baggage screeners may fail to detect (miss) an IED signal when employing a conservative β. Correspondingly, the signal may be missed because the resolution of the detection process is low in discriminating signals from noise, even if β is neutral or risky. Thus, an x-ray system that yields a high Pd is more sensitive than an x-ray system that fails to produce a signal which is not obscured by static and noise.

Sensitivity is a measure of the difference in average operator response levels as a function of the presence or absence of a signal. The perceptual analog of sensitivity, d', corresponds to the separation of the means of signal and noise distributions (figure B-2). As the signal magnitude increases, the mean of the signal and noise distribution moves to the left. As the magnitude of the signal decreases, the mean of the signal and noise distribution moves to the left. In each case, the proportion of signals detected (the P_d) changes as the distance between the signal and noise distributions varies. According to Wickens, if the separation between the distributions is great, sensitivity is great and a given operator response is quite likely to be generated by either signal or noise but not both. Similarly, if the separation between signal and noise is small, d' measures will be low.

Table B-1 provides a representative table of Z values and ordinate values of the probability distribution related to hit and false alarm responses. A complete table of the ordinate values of the standard normal distribution will be used to calculate d' for the Test and Evaluation Report (TER).

The procedures required to calculate d' are listed below (Coren and Ward).

Procedures to Calculate d'.

a. Find the false alarm rate from the outcome matrix in the HIT/FA column of table A-1;
b. Read across the table to the Z column (the label of the abscissa of the graph);
c. Calculate the value tabled there ORD(FA);
d. Repeat these operations for the hit rate, calling the tabled value Z(HIT);
e. Calculate d' using the following equation: d' = Z(FA) - Z(HIT).

TABLE B-1. REPRESENTATIVE Z-SCORES AND ORDINATE VALUES OF THE NORMAL CURVE FOR DIFFERENT RESPONSE PROBABILITIES TO CALCULATE β AND d'

HIT/FA	Z	ORD		HIT/FA	Z	ORD
.01	2.33	0.03		.50	0.00	0.40
.02	2.05	0.05		.55	-0.12	0.40
.03	1.88	0.07		.60	-0.25	0.39
.04	1.75	0.09		.65	-0.38	0.37
.05	1.64	0.10		.70	-0.52	0.35
.08	1.40	0.15		.75	-0.67	0.32
.10	1.28	0.18		.80	-0.84	0.28
.13	1.13	0.21		.82	-0.92	0.26
.15	1.04	0.23		.85	-1.40	0.23
.18	0.92	0.26		.88	-1.18	0.20
20	0.84	0.28		.90	-1.28	0.18
.25	0.67	0.32		.92	-1.40	0.15
.30	0.52	0.35		.95	-1.64	0.10
.35	0.38	0.37		.96	-1.75	0.09
.40	0.25	0.39		.97	-1.88	0.07
.45	0.12	0.40		.98	-2.05	0.05
.50	0.00	0.40		.99	-2.33	0.03

APPENDIX C
LINESCAN-TnT OPERATION

LINESCAN-TnT OPERATION

The Linescan-TnT.

The Linescan-TnT system is a training tool developed by EG&G Astrophysics to help baggage screeners operate the E-Scan x-ray equipment and identify security threats. For the IEDSS OT&E, the TnT system will be used to create a simulated environment for the training of security operators who use black/white and enhanced x-ray equipment. The TnT training for the OT&E will not teach screeners how to operate the black/white or enhanced x-ray equipment, rather, screeners will *review* previously completed lessons on detecting IEDs to improve their IED detection performance.

TnT Operations.

The TnT system consists of a combination of equipment and software including a color monitor, black/white monitor, control panel, and trackball. The TnT will be located away from the x-ray equipment in an area of SFO as specified by United Airlines. Participating baggage screeners will receive selected training lessons and certification tests. Both black/white and enhanced operators will receive the same TnT training lessons.

Screeners will choose the appropriate lessons from the TnT Lessons menu. Upon completion of an individual Lesson, subjects will review the answers of the respective test. All tests will have a score of 100 percent before conducting the OT&E. Baggage screeners will receive a copy of the baggage screener instructions (see below) one day before operating the TnT. A copy of the instructions will be located on the TnT, and an FAA representative will monitor all screeners operating the TnT.

Baggage Screener Instructions.

The Linescan TnT will be used to train of security operators who use black/white and enhanced x-ray equipment. Screeners will also receive the following instructions before beginning the TnT system:

This training system is used to train baggage screeners to improve performance in detecting threatening objects. By taking the following training lessons, you will learn to identify threats using both black/white and enhanced x-ray equipment. You will review 7 training lessons and 7 completed tests identified in the tables below list. Please do not take lessons that are not included on the list.

If you have any questions at any time, please contact a FAA representative.

Getting Started.

If you have any questions about beginning the training lessons, turning the equipment on, or using available functions, please notify a FAA representative.

Please follow the provided procedures to begin the training:

Logon Procedures

a.   Use trackball to position arrow on visual display number pad.

b.   Using the visual display number pad, enter operator ID: **123**
c.   Press the **ENTER** button to enter operator ID.
d.   Using the visual display number pad, enter password: **123**
e.   Press the **ENTER** button to enter password.
f.   Press the **CLEAR** button if you make a mistake to restart the logon procedures.
g.   Press the **TIP** button on the TnT control panel.

Reviewing a Lesson.

After you have successfully logged on to the training system:

a.   Select the **Lessons** button from the opening screen.
     The Lessons menu appears.
b.   On the Lessons menu, click **Review**.
     The Review menu appears.
c.   To review a prior lesson, click **A Prior Lesson**.

A checklist of the lessons previously completed is displayed. It shows up to 10 lessons at a time. To bring other lessons into view, use the up and down arrows to the right of the lessons.

     To select the lesson you want to review, click on it.

Table 1 shows a list of the lessons for you to review. Review the lessons in the order as presented in table 1.

List of Lessons to Review (table 1).

Unit 2,	Lesson 1
Unit 4,	Lesson 1
Unit 4	Lesson 3
Unit 5	Lesson 4
Unit 6	Lesson 1
Unit 6	Lesson 3
Unit 7	Lesson 1

When the lesson is presented to you, complete the lesson by closely reading the information presented on the screens.

a.   Begin the appropriate lesson displayed on right visual display (Lesson Screen).

b.   When finished reviewing a lesson, *immediately review the corresponding test.*
Review a Test.

Reviewing a test allows you to view a previously completed test. All tests you will review provide you
with the correct answer (highlighted in green) among four possible choices.

To review a test:

a.   On the Lessons menu, click **Review**.
          The Review menu appears.
b.   To review a prior test, click **A Prior Test**.

A checklist of the previously completed tests is displayed. It shows up to 10 tests at a time. To bring the
appropriate tests in to view, if needed, use the up and down arrows to the right of the tests.

          To select the test you want to review, click on it.

The appropriate tests for you to review are listed below in table 2. The correct answers of each test are
shown in green.

List of Tests to Review (table 2).

Unit 2,	Lesson 1
Unit 4,	Lesson 1
Unit 4	Lesson 3
Unit 5	Lesson 4
Unit 6	Lesson 1
Unit 6	Lesson 3
Unit 7	Lesson 1

a.   Review all questions and answers *closely*.

b.   To go to the next question, click the right arrow. To go to the previous question, click the left
          arrow.

When you have finished reviewing the test you are returned to the Lessons menu.

a.   Complete all lessons and tests as indicated in the above lists.
b.   When finished, select the **Exit** button until you reach the **Exit TnT** button.
c.   Press the **Exit TnT** button.